DOCUMENT RESUME

ED 338 044 FL 019 705

AUTHOR Davis, Katharine
TITLE The Acquisition of VOT: Is It Language-Dependent?
PUB DATE Jul 90
NOTE 11p.; In: Papers and Reports on Child Language Development, Number Twenty-nine. California, Stanford University, 1990. p28-37.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Articulation (Speech); Contrastive Linguistics; *English; *Hindi; *Language Acquisition; Language Research; Linguistic Theory; *Phonetics; Uncommonly Taught Languages
IDENTIFIERS *Voice Onset Time

ABSTRACT

A study investigated the acquisition of the phonemic feature/voice/ and, more specifically, voice onset time (VOT), in Hindi and English. Subjects were 20 monolingual children aged 2 to 6 from each of the 2 language groups, divided into 5 age groups, and 10 adults from each language group. Flash cards depicting minimally contrasting sets of words, differing only in the voicing of initial velar stop, were presented to each subject. The goal was to obtain five repetitions of each token from each speaker. Data were digitized and subjected to computer analysis. Lead and lag times were measured for each repetition of each token. The English data for post-release aspiration times confirmed earlier researchers' findings. Child and adult mean times did not differ substantially. The children had acquired a permanent voicing distinction by age 2. The Hindi children's mean values were not as adult-like as those of English-learners. It is concluded that there may be a universal pattern of VOT development that depends on the extent of differences in mean VOT times between adult target phonemes, and apparent age differences in phonemic distinctions may actually result from the languages' different phonetic categories. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

PRCLD 29 (1990)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED338044

THE ACQUISITION OF VOT: Is it language-dependent?

Katharine Davis
Cornell University

Introduction

Early theories of phonological acquisition concentrated on the phoneme as a whole. The criterion for determining that a phoneme had been acquired was the attainment of an adult-like production as judged by the listener. While researchers investigating these models did not particularly concern themselves with acoustic measurement, an adult-like production assumes that all the acoustic values associated with the sound in the adult model are present and correct. By definition then, when a phonemic contrast is evidenced, the acoustics have already been learned. There is no way to determine what stages, if any, the child goes through in his attainment of the correct production values. A more revealing approach from this perspective would be to break the phonemes down into their appropriate features and then to investigate the corresponding acoustic cues.

The present study is concerned with the acquisition of the phonemic feature /voice/, which is the distinguishing feature in the initial sounds of the pair, "cap" and "gap", for example. There is a general concensus in the literature that voice onset time, or VOT, (the amount of time, usually measured in milliseconds, between the onset of periodic vocal fold vibration and the release of oral closure), is the most salient perceptual cue to phonemic stop voicing.

The following schematic diagram (figure 1) illustrates how the phonemic systems of the languages to be discussed in this paper employ phonetic categories along the VOT time line.

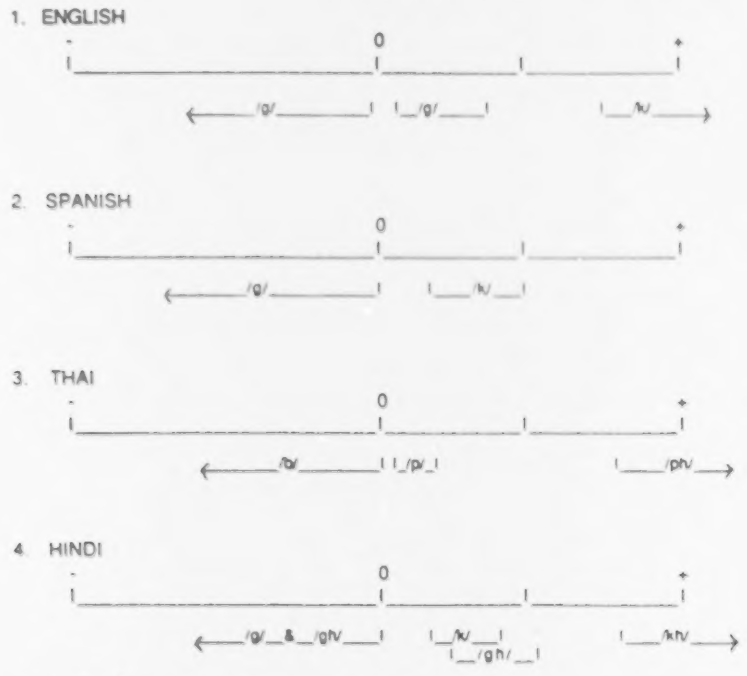

figure 1. Phonemic Systems - initial position (schematic)

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Eve Clark
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FL019705

BEST COPY AVAILABLE

Note that in English initial position the voiceless stop phoneme is phonetically realized as a long lag production. The voiced stop is realized with either prevoicing or short lag timing, depending to a large extent on the individual speaker. In Spanish, the voiceless stop has short lag timing, and the voiced stop must be prevoiced. Thai has three contrasting phonemes (except at the velar place), realized as indicated. The Hindi system is rather special in that the voiced aspirated stop has a fair amount of lag time in addition to the required prevoicing. With this background we may move on to the examination of some actual data.

In the mid-seventies, several studies on the acquisition of English voice onset time, (or VOT), were carried out, among them, Macken and Barton (1977), Zlatin and Koenigsnecht (1976), and Kewley-Port and Preston (1974). In spite of the fact that some of the studies were longitudinal and some were synchronic, they all agreed that in the very early stages of acquisition there seemed to be a concentration of stops produced in the short-lag VOT region.

By about the age of two, however, most children had acquired the distinction between voiced and voiceless phonemes as measured by significant mean differences of VOT. These studies also showed development toward the adult model over time. Although the details of this developmental course differed from study to study, both the narrowing of the range of within-phoneme-category VOT productions and movement toward the adult mean VOT value in each category were found to be part of the acquisition process in English.

Macken and Barton (1980) also had a follow-up study that looked at acquisition data from Spanish-speaking children. An interesting finding of theirs was that, while Spanish and English both have two-way voice contrasts in their phonological systems, (meaning that there are two and only two stops which differ by phonemic voice at each place of articulation), there was a major discrepancy in the ages at which children learned to reliably separate the voiced and voiceless stops in production.

The English-learning children had acquired the difference by the age of two, whereas the Spanish-learning children did not show production of this contrast even at the age of four. Macken and Barton suggested that this might be due to a spirantization rule in Spanish. That is to say, the children had actually learned spirantization as a way of making the phonemic distinction. Their target voiced phonemes spirantized much more than did their target voiceless phonemes, although they did not yet show consistent spirantization according to the adult rule. Because the children's rule was not yet fully developed, the stop and spirant production targets were confounded, and the children took longer to learn the phonetic values for the clear stops.

Following this argument, it would seem that children are required to have a knowledge of the rules of the phonological system as a whole before they may acquire the phonetic contrasts within a single feature.

Studies of additional languages have been carried out since then. In 1986, Gandour et al did a study of VOT acquisition in Thai, which has the three-way phonemic voicing contrast of voiced, voiceless unaspirated, and voiceless aspirated depicted above. They showed that the Thai-learning three-year olds, (like the Spanish four-year olds), do not have significantly different production times for voiceless unaspirated and voiced stops, even though Thai has nothing like the Spanish spirantization rule to confound the data. The paradox of different ages of phonemic voicing acquisition in different languages would thus appear to be as yet unresolved.

The question then, was whether this discrepancy in VOT acquisition was in fact due to differences in the phonologies of the two languages, or whether there was perhaps a less language-dependent way of accounting for it. In order to investigate the latter possibility, two languages with different phonemic voicing systems were chosen for study.

The present study looks at the acquisition of voice in Hindi and in English. As shown on the above chart, Hindi has a four-way phonemic voicing contrast and English has a two-way contrast.

Methodology

In order to test the production of VOT, flash cards were created for each language which depicted minimally contrasting sets of words which differed only in the voicing of an initial velar stop. All tokens were monosyllabic and none contained open syllables. For example, in English, there was a card with a picture of a goat on it and a card with a picture of a coat on it. There were six different sets for each language. A variety of vowel qualities was tested; results from the four English vowels and the three Hindi vowels shown here are presented below (figure 2).

English picture cards		Hindi picture cards			
goat	coat	kaan "ear"	khaan "mine"	gaan "anthem"	ghaan "quantity"
gate	Kate	kʌn "kernel"	khʌn "floor"	gʌN "group"	ghʌn "cube"
girl	curl	koT "fort"	khoT "flaw"	goT "border"	ghoT "grind"
game	came				

figure 2. Stimuli reported in the present study

The cards were presented in random order to monolingual children ranging from 2 to 6 years of age in individual recording sessions. The children were usually accompanied at the session by their mothers, or occasionally by another adult female (relative or caretaker), who were also recorded for use in determining the adult model.

The subjects were divided into groups according to their chronological age at the time of recording. For the two languages, there were five subjects in each of the four child groups (ie. 2-year-olds, 3-year-olds, 4-year-olds, and 5- and 6-year-olds). There were also ten adults selected at random from each language.

The goal was to obtain five repetitions of each token from each speaker. By and large we were successful. Occasionally, though, a child would object to a certain flashcard and refuse to give more than one or two repetitions, and sometimes a child would enthusiastically give eight or nine repetitions. All repetitions were included in the analysis. Vowel quality turned out not to interact significantly with VOT values for phonemic voice categories according to a three-way analysis of variance, (age, voice, and vowel), with repeated measures on voicing target and vowel.

All data were digitized and measured using the WAVES program at a Sun computer workstation. Spectrograms and waveforms were examined simultaneously in order to ensure an accurate identification of VOT durations.

Both lead and lag times were measured for each repetition of each token. This separation affords a better comparison of VOT values for both Hindi, in which the aspiration duration differences of the voiceless and voiced aspirates (/kh/ and /gh/) have been under question for many years, and in English, where the presence vs. absence of lead voicing is not phonemic. In

these cases, a total measurement of a token's prevoicing and post-release aspiration times together can obscure important facts about the data.

Voicing lead was measured from the start of periodic vocal fold vibration to the point of closure release. Post-release VOT was measured from the release of oral closure to the start of periodic vibration in the following vowel. Pre- and post-release measurements were entered separately in the analysis for the reasons just mentioned and to avoid underreporting the actual amount of lag time produced in the prevoiced tokens.

It is often claimed that the Hindi voiced aspirate has a distinctive production which requires loose, though periodic, vocal fold vibration (known as breathy voice or murmur) throughout the aspirated (post-release) portion of the consonant. This phenomenon was seen to occur in many of the productions, both child and adult, in this study. However, the large number of voiced aspirate repetitions with apparently normal aspiration may imply that voicing during aspiration is not, in fact, a requirement of the Hindi phonology, but rather an optional, (possibly style- or performance-related) production type. When murmur did occur, the post-release but pre-vocalic section was measured as normal lag time.

Aspiration amplitude and fundamental frequency measurements were also taken, in order to determine the acquisition patterns of redundant voicing cues, but only the VOT data are reported here.

Results

The English data for post-release aspiration times appear to confirm the findings of earlier researchers. As shown in the chart below (figure 3), the adult means were about 63ms for /k/ and about 20ms for /g/, with standard deviations of around 17 and 8ms respectively. The children's mean times did not differ substantially from the adults' except that their voiceless stops were a bit longer and their standard deviations were greater (showing less refinement toward the target).

Phoneme - Age Group	Count	Mean(ms)	Std. Dev.(ms)
k - 2	82	75	40
k - 3	154	74	40
k - 4	75	74	52
k - 6	117	74	21
k - Adult	194	63	17
g - 2	79	21	17
g - 3	169	19	13
g - 4	82	25	15
g - 6	111	22	12
g - Adult	211	20	8

figure 3. English Post-Release VOT Means by Age

A pairwise comparison of the VOT means is given below (figure 4). In this chart, the shaded area indicates that there is a significant difference of mean post-release voice onset times between the two phonemes. The numbers below the shaded area give the actual difference, the critical difference required, and the level of significance. The adult mean difference is also given at the left of the chart. The English speaking adults showed a mean production time for the voiceless stop which was about 43ms greater than the mean production time of the voiced stop, which was a highly significant difference. The fact that highly significant differences were evidenced for each age group indicates that the children in the English part of the study had acquired a permanent voicing distinction by the age of two, which confirms the findings of earlier studies.

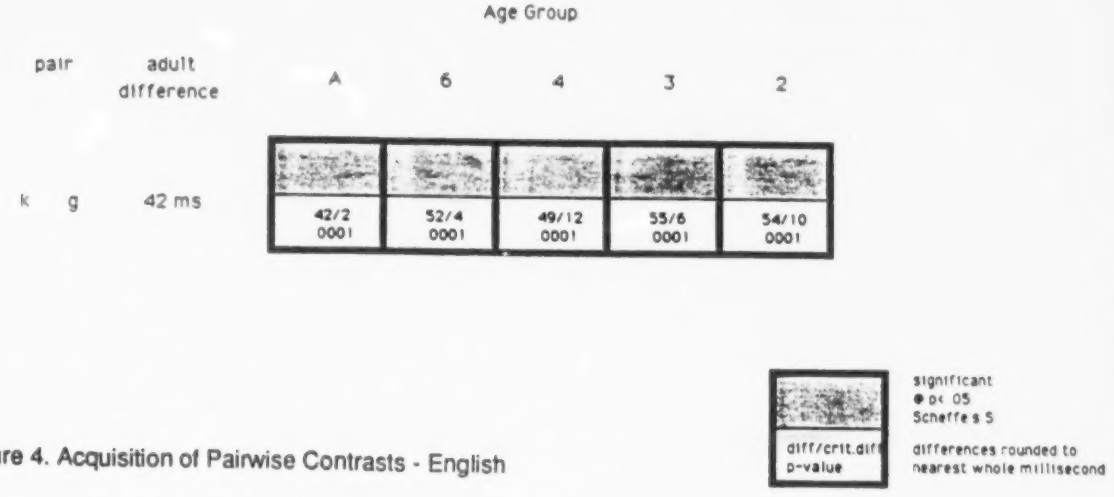

figure 4. Acquisition of Pairwise Contrasts - English

The table of post-release means and standard deviations for each Hindi phoneme by age group is given below (figure 5). Again, we found the children's standard deviations to be generally greater than the adults'. The Hindi children's mean values were not as adult-like as those of the English-learners.

Phoneme - Age Group	Count	Mean(ms)	Std. Dev (ms)
k - 2	77	21	18
k - 3	90	23	20
k - 4	86	18	13
k - 6	79	20	13
k - Adult	139	17	8
kh - 2	63	39	25
kh - 3	84	42	24
kh - 4	94	51	20
kh - 6	75	48	19
kh - Adult	158	51	14
g - 2	70	13	15
g - 3	93	21	22
g - 4	95	13	13
g - 6	80	8	10
g - Adult	156	9	10
gh - 2	73	30	23
gh - 3	92	32	26
gh - 4	93	34	24
gh - 6	68	25	24
gh - Adult	147	24	17

figure 5. Hindi Post-Release VOT Means by Age

The following chart of pairwise comparisons (figure 6) illustrates the nature of the adult-child differences noted in the interaction plot above. All pairs are significantly different in the adult group as indicated by the shaded areas in the first column. However, the younger groups do not exhibit acquisition of a contrast for all pairs in all age groups.

The 6-year-olds do not show a significant difference between members of the 1st pair (/gh/ vs. /k/); the 4-year-olds do not reliably distinguish the 2nd pair (/k/ vs. /g/); the 3-year-olds do not show significant mean differences for either of these two pairs, and the two-year-olds don't produce significant differences on either these pairs or the 4th pair (/kh/ vs. /gh/).

pair	adult difference	Age Group A	6	4	3	2
gh k	7ms	7/4 .0001	5/8 .3292	16/8 .0001	9/10 .0754	8/9 .1182
k g	9ms	9/4 .0001	13/8 .0001	4/8 .4304	3/9 .8884	9/10 .0665
gh g	15ms	15/4 .0001	16/8 .0001	20/7 .0001	12/9 .0082	17/10 .0001
kh gh	27ms	27/4 .0001	23/8 .0001	17/7 .0001	10/10 .0409	9/10 .0880
kh k	33ms	33/4 .0001	28/8 .0001	33/8 .0001	18/10 .0001	17/10 .0001
kh g	42ms	42/4 .0001	41/8 .0001	27/7 .0001	22/10 .0001	26/10 .0001

figure 6. Acquisition of Pairwise Contrasts - Hindi

Lead time data is presented, not as a function of mean differences between phoneme targets, but rather as a percentage of the voiced phoneme targets actually produced with lead voicing. Mean lead time differences would be unrevealing here due to the fact that all voiceless targets are assumed to have a lead voicing time of 0ms. (Although discussion is beyond the scope of this paper, it is interesting to note that in the one phonemic contrast in which both members are phonologically voiced, the adults did exhibit a significant difference of -60ms for the /g/ and -52ms for the /gh/.) The table below (figure 7) compares both English and Hindi data from the present study with data reported in Zlatin and Koenigsnecht (1976).

Age	Zlatin & Koenigsnecht English	English all tokens	English w/out abnormal productions	Hindi
2	4.5%	16%	2.5%	38%
6	8.5%	3%	3%	73%
A	34.3%	12%	12%	99%

figure 7. Percentage of Voiced Target Repetitions Produced with Lead Voicing

The 2-year-olds in the present English study showed some irregular productions of voicing lead. Some repetitions were prenasalized and some had a rather long (more than 20ms) gap of silence between the lead voicing and the closure release. (Allen,1985,reports a similar pre-nasalization phenomenon in the acquisition of lead voicing in French-learning children.) When the irregular productions were removed from the analysis, the percentage of repetitions with lead voicing were more in accord with the Zlatin and Koenigsnecht results. In both Hindi and English the percentage of voiced targets employing lead voicing increases with age.

Discussion

The lag time results of the present study are in accordance with both Macken and Barton studies. The Hindi contrast which is most similar phonetically to the Spanish contrast (i.e. the /k/ vs /g/ pair) has not been acquired at the age of four, but the contrast which is most similar phonetically to the English contrast (i.e. the /k/ vs /kh/ pair) is present even in the two-year old group.

Relying on phonological evidence alone to account for the differences in age of acquisition may make the development process appear more difficult than it needs to be. Macken and Barton's spirantization argument is essentially phonological in nature. By concentrating more on the phonetic contrasts involved in the phonemic distinction, a more straightforward picture emerges.

In other words, when the mean VOT difference between phoneme categories is great, the distinction is acquired quite early, as it is in English and in the fifth Hindi pair. When the mean VOT difference is small, the distinction takes longer to acheive, as in Spanish and in the second Hindi pair.

This explanation would also account for data in the Gandour et al study which shows, among other things, that /p/ and /ph/ are distinguished before /b/ and /p/. The mean VOT difference is much greater in the former pair.

Such a phonetic explanation does not imply that there is no involvement of the phonological system as a whole in the acquisition process. Other factors may interact with the process of learning phonological distinctions through phonetic cues.

For example, note that the two year-olds seem to have acquired a contrast on the third pair, (the voiced aspirate and the voiced unaspirate) before they have acquired the voiced aspirate vs voiceless aspirate distinction (pair four). This is in spite of the fact that the former pair has a mean VOT difference of 12ms less than the latter pair.

Obviously, if the children were using voice onset time differences alone, according to the present theory the reverse order of acquisition should have been the case. A possible explanation for this departure may lie in an across-the-board method of acquisition. Once a distinction has been learned for one contrast, it applies to all other appropriate cases.

Therefore, if the children have acquired the voiceless aspirate vs the voiceless unaspirate distinction early on, due to the large mean VOT difference of the adult model, they may have already concluded that aspiration is phonemic in Hindi. This would allow them to notice and produce the slightly smaller VOT differences in the voiced aspirate/unaspirate pair as well.

It is more difficult to explain why the four-year olds have a contrast while the six year-olds do not on the first pair, and vice versa on the second pair. It should be pointed out, though, that the mean VOT differences of the two pairs are very similar, and it may be due to chance that we see an apparent age-group difference here.

Note that these data are for post-release times only. So, this is not to say that the younger children are lacking a VOT contrast completely in some pairs. They may have indeed noticed that some of these sounds have prevoicing and some of them don't, and may be making some attempts at lead voicing. However, they are not forced to concern themselves with prevoicing at this stage, since reliable post-release differences do exist in the adult model. The presence of prevoicing would, of course, be even less important to the English-learning children since in English, adult voiced stops do not require prevoicing.

The small amount of lead voicing produced by children as compared to adults in both languages (and also in Gandour et al's Thai study), may be a result of its difficulty. Kewley-Port and Preston cite a great deal of articulatory evidence in their 1974 study which supports the argument that lead voicing is acquired last among VOT production types because of the physiological difficulty of producing it. It is reasonable to suppose that children may leave refinement of the more difficult lead voicing until later in the acquisition process.

Conclusion

In view of the acquisition data from the four-way Hindi system, we can posit that there may be a universal pattern of VOT development which depends on the extent of differences in mean VOT times between the adult target phonemes. Children are probably able to differentiate phoneme pairs with extreme VOT differences at an early age. It is only much later that finer differences can be produced accurately.

Therefore, apparent differences in the ages at which similar phonemic distinctions are made across languages may actually be the result of the different phonetic categories employed by those languages.

BIBLIOGRAPHY

Allen, George (1985). How the young French child avoids the pre-voicing problem for word-initial voiced stops. Journal of Child Language 12, 37-46.

Gandour, Jack and S.H. Petty, R. Dardarananda, S. Dechongkit, S. Mukngoen (1986). The acquisition of the voicing contrast in Thai: A study of voice onset time in word-initial stop consonants. Journal of Child Language 13, 561-572.

Kewley-Port, Diane and M.S. Preston (1974). Early apical stop production: A voice onset time analysis. Journal of Phonetics 2, 195-210.

Macken, Marlys A. and D. Barton (1977). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. Journal of Child Language 7, 41-74.

---(1980). The acquisition of the voicing contrast in Spanish: A phonetic and phonological study of word-initial stop consonants. Journal of Child Language 7, 433-458.

Zlatin, Marsha A. and R. A. Koenigsnecht (1976). Development of the voicing contrast: A comparison of voice onset time in perception and production. Journal of Speech and Hearing Research 19, 78-92.

I would like to thank Raj Bordia, Toni Borowsky, Walter Cichocki, John Kingston, and Barbara Lust for useful discussion of these results.